MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2018/2019

DPS5018 –INTRODUCTION TO PROBABILITY AND STATISTICS

(For Diploma students only)

8 MARCH 2019
9.00 a.m - 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **6** pages (**2** pages with **4** questions and **4** pages for appendix)
2. Answer **ALL** questions. All necessary working steps must be shown.
3. Write all your answers in the answer booklet provided.
4. Statistical Tables will be provided.

QUESTION 1 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

(a) The following table shows the average times spent (minutes) on studying at night by 30 randomly chosen students in MMU.

30	40	88	35	32	47	62	35	59	41
54	69	58	72	70	73	60	33	82	80
44	50	46	71	33	80	36	82	40	32

Construct a group data table using the classes 30–39, 40–49, and so on. [4 marks]

(b) A random sample is taken of 10 independent measurements for body temperature, in degree Celsius. The values obtained are given below. Find the sample standard deviation. [6 marks]

37.1	36.5	36.8	37.0	36.9	36.5	37.2	36.2	36.4	36.6

QUESTION 2 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

(a) The following probability distribution recorded the number of visits by shoppers to a convenience store per week.

Number of visits, x	1	2	3	4	5
P $(X=x)$	0.16	0.22	0.33	0.16	0.13

i. Find $P(X \leq 3)$. [2 marks]

ii. Find the mean and variance of this distribution. [5 marks]

(b) From the previous result, the probability that a student will get an 'A' in the coursework marks of Essential Statistics is 0.65. If a sample of 10 students is selected at random, find the probability that exactly 6 students will get an 'A' for the coursework marks. [3 marks]

Continued...

QUESTION 3 [10 MARKS]
[CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]

(a) A random sample of 539 households from a certain city was selected and it was found that 133 of these households own at least one fire extinguisher.

i. Find the point estimate of the population proportion. [1 mark]

ii. Construct a 95% confidence interval for the population proportion. [2.5 marks]

(b) A company claims that its light bulbs are better than those of its main competitor. If a study shows that a sample of 40 of its bulbs has a mean lifetime of 647 hours of continuous use, while a sample of 40 bulbs that made by the main competitor has a mean lifetime of 638 hours of continuous use. Assume that the standard deviation for the bulbs made by the company and its main competitor are 27 hours and 31 hours respectively.

i. State the null and alternative hypothesis. [1 mark]

ii. By using the hypothesis testing of two population mean, calculate the value of the test statistic. [1.5 marks]

iii. At the 10% significance level, do the data provide sufficient evidence to support the company's claim? [4 marks]

QUESTION 4 [10 MARKS]
[FOR (b), CORRECT YOUR ANSWERS TO FOUR DECIMAL PLACES.]
It is believed that there is a linear relationship between house price (in thousand dollar), y, and its square feet, x. The following data show houses' square feet and their prices.

Square Feet, x	1100	1400	1425	1550	1600	1700	1725	1875	2350	2450
House Price (in thousand dollar), y	199	245	319	219	312	255	279	308	405	324

(a) Construct a table which contains x, y, x^2, y^2, xy and their individual sums. [2 marks]

(b) Find

i. the values of S_{XX}, S_{YY}, S_{XY}, y-intercept β_0 and slope β_1. [5 marks]

ii. the simple linear regression that fits the data obtained from (i). [0.5 mark]

iii. the house price when the house square feet is 2000. [0.5 mark]

(c) Compute and interpret the sample correlation coefficient, r. [2 marks]

Continued...

APPENDIX – KEY FORMULA

- **Mean for ungrouped data**

$$\bar{x} = \frac{\sum x_i}{n} \quad \text{or} \quad \mu = \frac{\sum x_i}{N}$$

- **Mean for grouped data**

$$\bar{x} = \frac{\sum x_i f_i}{n} \quad \text{or} \quad \mu = \frac{\sum x_i f_i}{N}$$

Where x is the midpoint and f is the frequency of the class

- **Variance for ungrouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad \text{or} \quad s^2 = \frac{\sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n}}{n-1}$$

- **Variance for grouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{n-1} \quad \text{or} \quad s^2 = \frac{\sum x_i^{\,2} f_i - \frac{\left(\sum x_i f_i\right)^2}{n}}{n-1}$$

Where x is the midpoint and f is the frequency of the class

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

- Mean of a discrete random variable x : $\mu = \sum xP(X = x)$
- Standard deviation of a discrete random variable x : $\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$
- Binomial probability formula : $P(X = x) = \binom{n}{x} p^x q^{n-x}$
- Poisson probability formula : $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$

Continued...

- Key formula for Binomial Probability Distribution:
  - If $p \leq 0.5$
    i. $P(X = r) = B(r) - B(r-1)$
    ii. $P(X \geq r) = 1 - B(r-1)$
    iii. $P(X \leq r) = B(r)$
    iv. $P(X > r) = 1 - B(r)$
    v. $P(X < r) = B(r-1)$
    vi. $P(a \leq X \leq b) = B(b) - B(a-1)$

  - If $p > 0.5$
    i. $P(X = r) = B(n-r) - B(n-r-1)$
    ii. $P(X \geq r) = B(n-r)$
    iii. $P(X \leq r) = 1 - B(n-r-1)$
    iv. $P(a \leq X \leq b) = B(n-a) - B(n-b-1)$

- Key formula for Poisson Probability Distribution:
  - If $\mu = \lambda$
    i. $P(X = r) = Poi(r) - Poi(r-1)$
    ii. $P(X \geq r) = 1 - Poi(r-1)$
    iii. $P(X \leq r) = Poi(r)$
    iv. $P(X > r) = 1 - Poi(r)$
    v. $P(X < r) = Poi(r-1)$
    vi. $P(a \leq X \leq b) = Poi(b) - Poi(a-1)$

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \dfrac{x - \mu}{\sigma}$

Continued...

ESTIMATION

- The $(1-\alpha)100\%$ confidence interval for μ is

$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

$\bar{x} \pm t_{\alpha/2,n-1}s_{\bar{x}}$ where $s_{\bar{x}} = \dfrac{s}{\sqrt{n}}$

- The $(1-\alpha)100\%$ confidence interval for p is

$\hat{p} \pm z_{\alpha/2}s_{\hat{p}}$ where $s_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

- **Sample size**

$$n = \left[\frac{\left(Z_{\alpha/2}\right)\sigma}{E}\right]^2 \qquad n = \frac{\left(Z_{\alpha/2}\right)^2 pq}{E^2}$$

HYPOTHESIS TESTING

- **ONE population**

Mean:

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \textbf{or} \quad t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

Proportion:

$$Z = \frac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$$

- **TWO population**

Mean:

$$Z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^{\,2}}{n_1}+\dfrac{\sigma_2^{\,2}}{n_2}}}$$

Proportion:

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \qquad \text{where} \qquad \hat{p}_p = \frac{x_1+x_2}{n_1+n_2}$$

Continued...

REGRESSION

- For the **least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$

  where $\beta_1 = \dfrac{S_{xy}}{S_{xx}}$ and $\beta_0 = \dfrac{1}{n}\left[\sum y - \beta_1 \sum x\right]$

- **Correlation coefficient,** r

  $$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

  where

  $$S_{xx} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}, \qquad S_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{n} \text{ and}$$

  $$S_{xy} = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

- **Error of sum of squares, SSE**

  $$SSE = S_{yy} - \frac{\left(S_{xy}\right)^2}{S_{xx}}$$

- **Standard error of the estimate,** s_e

  $$s_e = \sqrt{\frac{SSE}{n-2}}$$

- **Regression** t**-Test**

  $$t = \frac{b_1 - \beta_1}{s_b} \qquad \text{where } s_b = \frac{s_e}{\sqrt{S_{xx}}} \qquad \text{and } s_e = \sqrt{\frac{S_{yy} - b_1 S_{xy}}{n-2}}$$

- **Correlation** t**-Test**

  $$t = \frac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}$$

End of Page.